Yulin Wang

Ph.D. Student, Tsinghua University

☎ +86 131-2121-5969 • ✉ wang-yl19@mails.tsinghua.edu.cn

🌐 www.wyl.cool

Education

Department of Automation, Tsinghua University

Ph.D. Student in Machine Learning and Computer Vision   *2019.09 – 2025.04 (expected)*

- Advisors: Prof. Cheng Wu and Prof. Gao Huang.

School of Automation Science and Electrical Engineering, Beihang University

B.Eng. in Automation   *2015.09 – 2019.06*

- GPA *Top 1/231*.
- Recipient of the **"Shen Yuan" Medal** (the highest honor for undergraduate students, **10/18,000+**).

Research Experience

Berkeley Deep Drive, University of California, Berkeley

Research Intern (advisor: Dr. Ching-Yao Chan)   *2018.07 – 2018.09*

Lab of Intelligent Manufacturing, Beihang University

Research Intern (advisor: Prof. Fei Tao)   *2017.06 – 2018.06*

Research Interests

Yulin Wang's research has centered on addressing challenges in *computational efficiency* and *data efficiency* when building and deploying large-scale foundation AI models (*e.g.*, large vision/multimodal models, generative models, and embodied foundation models for robotics). His goal is to enable high-level intelligence at an affordable training and inference cost.

During his Ph.D., he discovered that introducing human-like *adaptiveness* into deep learning could be an effective approach to achieving this goal. Specifically, he took inspiration from human cognition processes and neuroscience, and developed algorithms capable of *adaptively adjusting inference and learning strategies conditioned on the diversified information and patterns within different data*.

Publications

Yulin Wang has published a number of works in top-tier conferences & journals in the fields of machine learning and computer vision, including *TPAMI* (**6**), *IJCV* (**2**), **NeurIPS** (**5**), **ICLR** (**1**), **ICCV** (**6**), **CVPR** (**5**), and **ECCV** (**2**). He has collected *more than 2,500 citations* according to **Google Scholar**.

I. Selected Publications

I.1 Adaptive inference for large-scale perception/generative models

> **TL;DR:** *Dynamically adjusting inference strategies based on each of the diversified input/generated samples (e.g., mimicking human vision, modeling visual perception as a sequential decision process, sequentially and actively deciding 'where to see' and 'when to conclude observation'), thus improving inference efficiency.*

[1] - **Yulin Wang**, Haoji Zhang, Yang Yue, Shiji Song, Chao Deng, Junlan Feng, Gao Huang
   Uni-AdaFocus: Spatial-temporal Dynamic Computation for Video Recognition
   IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI, IF$_{5\text{-year}}$=22.2), 2024

   - **Yulin Wang**, Zhaoxi Chen, Haojun Jiang, Shiji Song, Yizeng Han, Gao Huang
   Adaptive Focus for Efficient Video Recognition
   IEEE/CVF International Conference on Computer Vision (ICCV Oral, acceptance rate: 3%), 2021

[2] - Gao Huang*, **Yulin Wang***, Kangchen Lv, Haojun Jiang, Wenhui Huang, Pengfei Qi, Shiji Song
[*co-first author with my advisor*]
Glance and Focus Networks for Dynamic Visual Recognition
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI, IF_{5-year}=22.2), 2023
- **Yulin Wang**, Kangchen Lv, Rui Huang, Shiji Song, Le Yang, Gao Huang
Glance and Focus: A Dynamic Approach to Reducing Spatial Redundancy in Image Classification
Advances in Neural Information Processing Systems (NeurIPS), 2020

[3] **Yulin Wang**, Rui Huang, Shiji Song, Zeyi Huang, Gao Huang
Not All Images are Worth 16x16 Words: Dynamic Transformers for Efficient Image Recognition
Advances in Neural Information Processing Systems (NeurIPS), 2021

I.2 Efficient training of foundation models

TL;DR: Designing tailored learning strategies for each type of discriminative pattern within data (e.g., low/high-frequency patterns and local/global patterns in vision), and organizing them into a unified training procedure mimicking human learning (e.g., through curriculum learning or neuroscience-inspired mechanisms), thus training large foundation models more efficiently, stably, and effectively.

[4] - **Yulin Wang**, Yang Yue, Rui Lu, Yizeng Han, Shiji Song, Gao Huang
EfficientTrain++: Generalized Curriculum Learning for Efficient Visual Backbone Training
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI, IF_{5-year}=22.2), 2024
- **Yulin Wang**, Yang Yue, Rui Lu, Tianjiao Liu, Zhao Zhong, Shiji Song, Gao Huang
EfficientTrain: Exploring Generalized Curriculum Learning for Training Visual Backbones
IEEE/CVF International Conference on Computer Vision (ICCV), 2023

[5] - **Yulin Wang**, Zanlin Ni, Yifan Pu, Cai Zhou, Jixuan Ying, Shiji Song, Gao Huang
InfoPro: Locally Supervised Deep Learning by Maximizing Information Propagation
International Journal of Computer Vision (IJCV, IF_{5-year}=14.5), 2024
- **Yulin Wang**, Zanlin Ni, Shiji Song, Le Yang, Gao Huang
Revisiting Locally Supervised Learning: An Alternative to End-to-end Training
International Conference on Learning Representations (ICLR), 2021

I.3 Data-efficient learning/fine-tuning

TL;DR: Mimicking "learning by analogy", augmenting the semantic diversity of training samples based on their individual characteristics, thus reducing the expensive cost of collecting high-quality training data or fine-tuning data (e.g., for supervised fine-tuning and transfer learning).

[6] - **Yulin Wang**, Gao Huang, Shiji Song, Xuran Pan, Yitong Xia, Cheng Wu
Regularizing Deep Networks with Semantic Data Augmentation
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI, IF_{5-year}=22.2), 2022
- **Yulin Wang**, Xuran Pan, Shiji Song, Hong Zhang, Cheng Wu, Gao Huang
Implicit Semantic Data Augmentation for Deep Networks
Advances in Neural Information Processing Systems (NeurIPS), 2019

II. All Publications

II.1 All Publications 1/3 – Adaptive inference for large-scale perception/generative models

[7] **Yulin Wang**, Yang Yue, Xinhong Xu, Ali Hassani, Victor Kulikov, Nikita Orlov, Shiji Song, Humphrey Shi, Gao Huang
AdaFocus V3: On Unified Spatial-temporal Dynamic Video Recognition
European Conference on Computer Vision (ECCV), 2022

[8] **Yulin Wang**, Yang Yue, Yuanze Lin, Haojun Jiang, Zihang Lai, Victor Kulikov, Nikita Orlov, Humphrey Shi, Gao Huang
AdaFocus V2: End-to-End Training of Spatial Dynamic Networks for Video Recognition
*IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2022*

[9] Yang Yue*, **Yulin Wang***, Bingyi Kang, Yizeng Han, Shenzhi Wang, Shiji Song, Jiashi Feng, Gao Huang　　[*co-first author, supervising Yang Yue (junior Ph.D. student)]*
Dynamic Inference of Multimodal Large Language Models for Efficient Robot Execution
*Advances in Neural Information Processing Systems (**NeurIPS**), 2024*

[10] Zanlin Ni*, **Yulin Wang***, Renping Zhou, Yizeng Han, Jiayi Guo, Zhiyuan Liu, Yuan Yao, Gao Huang　　[*co-first author, supervising Zanlin Ni (junior Ph.D. student)]*
ENAT: Rethinking Spatial-temporal Interactions in Token-based Image Synthesis
*Advances in Neural Information Processing Systems (**NeurIPS**), 2024*

[11] Zanlin Ni*, **Yulin Wang***, Renping Zhou, Rui Lu, Jiayi Guo, Jinyi Hu, Zhiyuan Liu, Yuan Yao, Gao Huang　　[*co-first author, supervising Zanlin Ni (junior Ph.D. student)]*
AdaNAT: Exploring Adaptive Policy for Token-Based Image Generation
*European Conference on Computer Vision (**ECCV**), 2024*

[12] Zanlin Ni*, **Yulin Wang***, Renping Zhou, Jiayi Guo, Jinyi Hu, Zhiyuan Liu, Shiji Song, Yuan Yao, Gao Huang　　[*co-first author, supervising Zanlin Ni (junior Ph.D. student)]*
Revisiting Non-Autoregressive Transformers for Efficient Image Synthesis
*IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2024*

[13] Ziwei Zheng, Le Yang, **Yulin Wang**, Miao Zhang, Lijun He, Gao Huang, Fan Li
Dynamic Spatial Focus for Efficient Compressed Video Action Recognition
*IEEE Transactions on Circuits and Systems for Video Technology (**TCSVT, IF**$_{5\text{-year}}$**=7.1**), 2023*

[14] Yizeng Han, Gao Huang, Shiji Song, Le Yang, Honghui Wang, **Yulin Wang**
Dynamic Neural Networks: A Survey
*IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI, IF**$_{5\text{-year}}$**=22.2**), 2022*

[15] Ziwei Zheng, Zechuan Zhang, **Yulin Wang**, Shiji Song, Gao Huang, Le Yang
Rethinking the Architecture Design for Efficient Generic Event Boundary Detection
*ACM International Conference on Multimedia (**ACM MM**), 2024*

[16] Yizeng Han, Dongchen Han, Zeyu Liu, **Yulin Wang**, Xuran Pan, Yifan Pu, Chao Deng, Junlan Feng, Shiji Song, Gao Huang
Dynamic Perceiver for Efficient Visual Recognition
*IEEE/CVF International Conference on Computer Vision (**ICCV**), 2023*

[17] Yifan Pu, Yiru Wang, Zhuofan Xia, Yizeng Han, **Yulin Wang**, Weihao Gan, Zidong Wang, Shiji Song, Gao Huang
Adaptive Rotated Convolution for Rotated Object Detection
*IEEE/CVF International Conference on Computer Vision (**ICCV**), 2023*

[18] Le Yang, Haojun Jiang, Ruojin Cai, **Yulin Wang**, Shiji Song, Gao Huang, Qi Tian
CondenseNet V2: Sparse Feature Reactivation for Deep Networks
*IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2021*

II.2 All Publications 2/3 – Efficient training of foundation models

[19] Zanlin Ni*, **Yulin Wang***, Jiangwei Yu, Haojun Jiang, Yue Cao, Gao Huang
[**co-first author, supervising Zanlin Ni (junior Ph.D. student)*]
Deep Incubation: Training Large Models by Divide-and-Conquering
*IEEE/CVF International Conference on Computer Vision (**ICCV**), 2023*

II.3 All Publications 3/3 – Data-efficient learning/fine-tuning

[20] Chaoqun Du, **Yulin Wang**, Shiji Song, Gao Huang
Probabilistic Contrastive Learning for Long-Tailed Visual Recognition
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI, IF$_{\text{5-year}}$=22.2), 2024

[21] **-** Mixue Xie, Shuang Li, Kaixiong Gong, **Yulin Wang**, Gao Huang
Adapting Across Domains via Target-Oriented Transferable Semantic Augmentation Under Prototype Constraint
International Journal of Computer Vision (IJCV, IF$_{\text{5-year}}$=14.5), 2023
- Shuang Li, Mixue Xie, Kaixiong Gong, Chi Harold Liu, **Yulin Wang**, Wei Li
Transferable Semantic Augmentation for Domain Adaptation
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR Oral), 2021

[22] Yifan Pu, Yizeng Han, **Yulin Wang**, Junlan Feng, Chao Deng, Gao Huang
Fine-grained Recognition with Learnable Semantic Data Augmentation
IEEE Transactions on Image Processing (TIP, IF$_{\text{5-year}}$=12.1), 2023

[23] Wenxuan Ma, Shuang Li, Jinming Zhang, Chi Harold Liu, Jingxuan Kang, **Yulin Wang**, Gao Huang
Borrowing Knowledge From Pre-trained Language Model: A New Data-efficient Visual Learning Paradigm
IEEE/CVF International Conference on Computer Vision (ICCV), 2023

[24] Wenxuan Ma, Jinming Zhang, Shuang Li, Chi Harold Liu, **Yulin Wang**, Wei Li
Making the Best of Both Worlds: A Domain-Oriented Transformer for Unsupervised Domain Adaptation
ACM International Conference on Multimedia (ACM MM), 2022

[25] Shuang Li, Kaixiong Gong, Chi Harold Liu, **Yulin Wang**, Feng Qiao, Xinjing Cheng
MetaSAug: Meta Semantic Augmentation for Long-Tailed Visual Recognition
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021

Selected Awards and Honors

○ National Scholarship for Ph.D. Students, Ministry of Education of China *2021, 2023*
  (2 times, Top 2% Ph.D. students in Tsinghua University)

○ **Microsoft Research Ph.D. Fellowship**, Microsoft Research Asia *2022*
  (12 Ph.D. students in the Asia-Pacific region)

○ **ByteDance Scholarship**, ByteDance Ltd. *2022*
  (10 Ph.D. students in China)

○ **Baidu Scholarship**, Baidu Inc. *2021*
  (10 Ph.D. students worldwide)

○ **CCF-CV Outstanding Young Researcher Award**, China Computer Federation (CCF) *2021*
  (3 Ph.D./MS students in China)

- Travel Award, NeurIPS _2019_
- **"Shen Yuan" Medal**, Beihang University _2018_
  (the highest honor for undergraduate students, 10/18,000+)
- National Scholarship for Undergraduates, Ministry of Education of China _2017, 2018_
  (**2 times, Top 2% undergraduates in Beihang University**)
- "Gong Xin" Innovation Scholarship, Ministry of Industry and Information Technology of China _2017_
  (**Top 1/231 in Beihang University**)
- First Prize, National Undergraduate Mathematical Contest in Modeling _2017_
  (**Top 0.2%**)

Teaching Experience

- Guest Lecturer for _Pattern Recognition and Machine Learning_ _Spring 2023, 2024_
- Teaching Assistant for _Nonlinear Programming_ _Fall 2021, 2022, 2023_

Academic Service

- Reviewer for TPAMI, IJCV, TCYB, TNNLS, TCSVT, Pattern Recognition, TMLR, ...
- Reviewer for ICML, NeurIPS, ICLR, CVPR, ICCV, ECCV, AAAI, ...
  - _Outstanding Reviewer_, CVPR, 2021
- Co-sponsor of the Special Interest Group on Dynamic Neural Networks, Beijing Academy of Artificial Intelligence (BAAI).
  - https://littlepure2333.github.io/dynamic-neural-network
  - Core members include more than 20 researchers from 8 universities. We have organized more than 30 academic reports and tutorials. The cumulative audience has exceeded 1,000.

Invited Talks and Presentations

- 2023.09, Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Dynamic Inference of Neural Networks
- 2023.02, School of Automation, Beijing Institute of Technology, Vision Transformers Meet Dynamic Inference
- 2021.12, PRCV 2021, Dynamic Deep Networks for Reducing Spatial Redundancy
- 2021.10, School of Computer Science, Fudan University, Dynamic Deep Networks for Reducing Spatial Redundancy
- 2021.09, Aibee (invited by Yuanqing Lin), Semantic Data Augmentation
- 2021.06, AI Time, Locally Supervised Deep Learning
- 2021.04, Beijing Academy of Artificial Intelligence, Dynamic Image/Video Recognition
- 2021.03, ByteDance Ltd., Semantic Data Augmentation
- 2020.11, Qingyuan Seminar, Glance and Focus Networks
- 2020.06, Huawei Technologies Ltd., Glance and Focus Networks
- 2019.10, School of Computer Science and Engineering, Beihang University, Semantic Data Augmentation